Opening of new catering distributor centers in Minsk, Belarus

1. Introduction/Business Problem

One distribution company decided to open its representative office in Minsk to sell various products for catering: kitchen and restaurant equipment, reusable and disposable tableware. In addition, it imports and sells premium coffee, tea and various snacks, etc. It is expected that about 5 warehouses will be opened in the city, from which goods will be delivered to the nearest points of sale.

Objective:

The objective of the study is to identify areas for each of the 5 points that will optimally correspond to the logistics tasks.

2. Data

- Foursquare API - to get the venue data from location
- geolocator librarys - to update zip for venues that not represented in Foursquare
- Folium library - to plot the geo data
- sklearn - To do clustering on the data
- panda, nampu - python libraries to manupulate with data

All these services/libraries has been used to fetch and analyze existing foodservice points in Minsk city where customer's firm can apply its business plans

2.1 Data sources

Foursquare service has been used as main data source for venues.

2.2 Data cleaning

Data downloaded or scraped from Foursquare sources were combined into one table. As a result we got a lot of non-needed and missing values(postal code, for example). Missed postal codes have been fetched with additional geolocator library and all remained NaN records have been skipped.

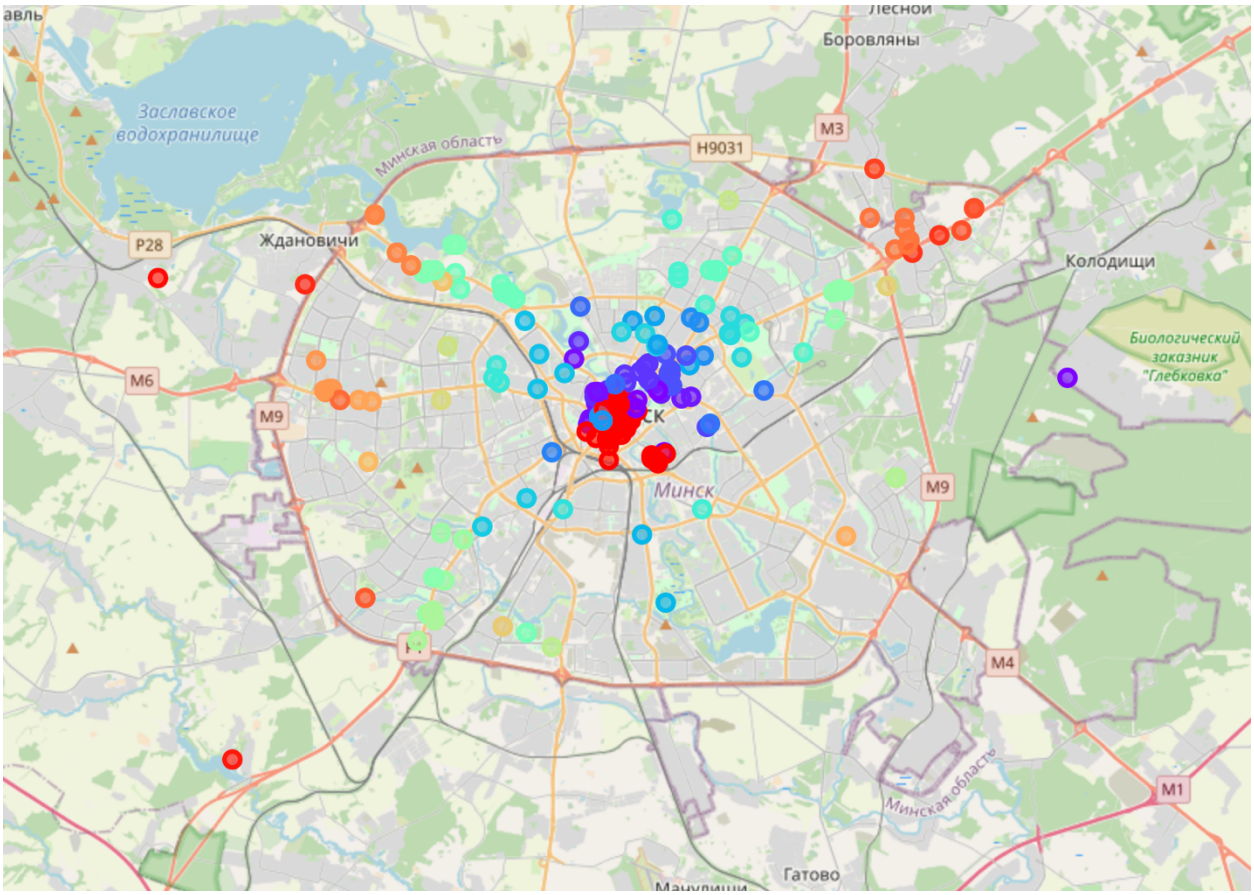
2.3 Feature selection

We worked only with the following attributes for selected venues:

'venue.name',
'venue.location.address',
'venue.location.lat',
'venue.location.lng',
'venue.location.distance',
'venue.location.formattedAddress',
'venue.categories',
'venue.location.postalCode'

3. Exploratory Data Analysis

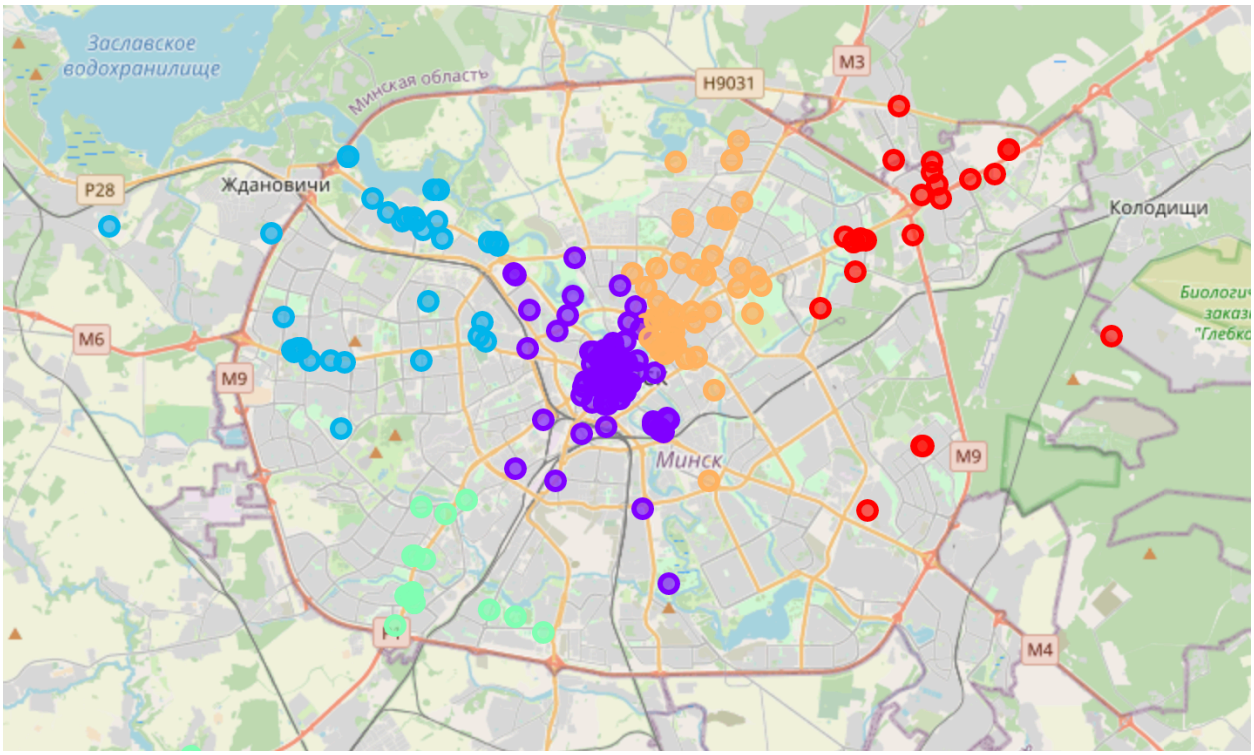
First off all we display all selected venues on the map to cleaner understand the pictures



Here we got an idea only for all the different postal codes.

4. Predictive Modeling

To split all venues on 5 areas applied k-means clustering



5. Conclusions

Received 5 equivalent areas in which the customer can place their centers. Given the specifics of the area / development / property prices, the customer now needs to find a suitable place in this area where his warehouse / office will be located